Lecture 30

In this lecture we continue our study of conditional probability, looking at Bayes' Theorem.

Recall that a <u>partition</u> of a set S is a way of dividing up S into disjoint nonempty subsets.

More formally, a partition of a set S is a collection of subsets

$$\{A_1, A_2, \ldots, A_n\}^{\dagger}$$

of S such that

$$S = A_1 \cup A_2 \cup \cdots \cup A_n$$

and the A_i are nonempty and pairwise disjoint.

Now let's suppose that S is a sample space. Let B be any event on S.

Then we have the following theorem.

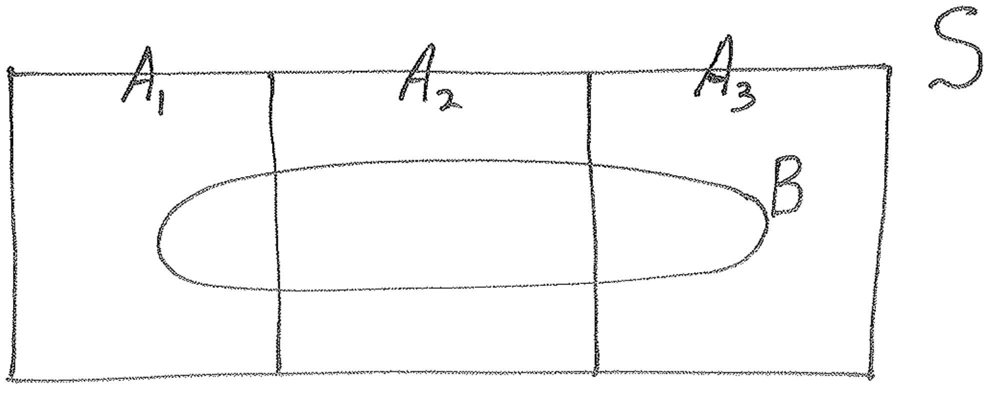
$\dagger$ In some situations n can be infinite.

Theorem

Let $\{A_1, A_2, \ldots, A_n\}$ be a partition of a sample space S. Then

$$\Pr(B) = \Pr(A_1 \cap B) + \Pr(A_2 \cap B) + \cdots + \Pr(A_n \cap B)$$
$$= \sum_{j=1}^{n} \Pr(A_j \cap B).$$

Here's an illustration of the situation when $n = 3$.



Now recall that conditional probabilities are defined using intersections:

$$\Pr(D|C) = \frac{\Pr(C \cap D)}{\Pr(C)}$$

Rearranging this gives:

$$Pr(C \cap D) = Pr(D|C) \cdot Pr(C)$$

We can apply this to the intersections given in the previous theorem:

$$Pr(A_j \cap B) = Pr(B|A_j) \cdot Pr(A_j)$$

So the theorem can be restated as follows:

$$\begin{aligned} Pr(B) &= Pr(B|A_1) \cdot Pr(A_1) + \cdots + Pr(B|A_n) \cdot Pr(A_n) \\ &= \sum_{j=1}^{n} Pr(B|A_j) \cdot Pr(A_j) \end{aligned}$$

Then for each $i = 1, 2, \ldots, n$ we have that

$$Pr(A_i|B) = \frac{Pr(A_i \cap B)}{Pr(B)}.$$

Putting everything together gives us the following theorem.

Bayes' Theorem

If $\{A_1, A_2, \ldots, A_n\}$ is a partition of a sample space S, and B is any event, then for each $i = 1, 2, \ldots, n$ we have that

$$Pr(A_i \mid B) = \frac{Pr(B|A_i) \cdot Pr(A_i)}{\sum_{j=1}^{n} Pr(B|A_j) \cdot Pr(A_j)} .$$

E.g. A die is cast.

$S = \{1, 2, 3, 4, 5, 6\}$

$A_1 = \{1, 2\}$, $A_2 = \{3, 4, 5\}$, $A_3 = \{6\}$

Let B be the event that "an even number is scored".

(i) Given that B occurs, what's the probability that the score is in A_1?

(ii) Repeat for A_2.

(iii) Repeat for A_3.

Sol$^{\underline{n}}$

Firstly we calculate the probability that each A_i occurs when a die is cast:

$$Pr(A_1) = \frac{2}{6} = \frac{1}{3}$$

$$Pr(A_2) = \frac{3}{6} = \frac{1}{2}$$

$$Pr(A_3) = \frac{1}{6}$$

Now we calculate the conditional probability of B given each A_i:

$$Pr(B|A_1) = \frac{Pr(A_1 \cap B)}{Pr(A_1)} = \frac{1}{2}$$

$$Pr(B|A_2) = \frac{Pr(A_2 \cap B)}{Pr(A_2)} = \frac{1}{3}$$

$$Pr(B|A_3) = \frac{Pr(A_3 \cap B)}{Pr(A_3)} = \frac{1}{1} = 1$$

So we can answer each part of the question:

(i)

$$Pr(A_1|B) = \frac{Pr(B|A_1)\cdot Pr(A_1)}{Pr(B|A_1)\cdot Pr(A_1) + Pr(B|A_2)\cdot Pr(A_2) + Pr(B|A_3)\cdot Pr(A_3)}$$

$$= \frac{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right) + (1)\left(\frac{1}{6}\right)}$$

$$= \frac{\left(\frac{1}{6}\right)}{\left(\frac{1}{6}\right) + \left(\frac{1}{6}\right) + \left(\frac{1}{6}\right)}$$

$$= \frac{\left(\frac{1}{6}\right)}{\left(\frac{3}{6}\right)} = \frac{1}{3}$$

(ii)

$$Pr(A_2|B) = \frac{Pr(B|A_2)\cdot Pr(A_2)}{Pr(B|A_1)\cdot Pr(A_1) + Pr(B|A_2)\cdot Pr(A_2) + Pr(B|A_3)\cdot Pr(A_3)}$$

$$= \frac{\left(\frac{1}{3}\right)\left(\frac{1}{2}\right)}{\left(\frac{3}{6}\right)} = \frac{\left(\frac{1}{6}\right)}{\left(\frac{3}{6}\right)} = \frac{1}{3}$$

(iii) $Pr(A_3|B) = ?$

The new numerator is $Pr(B|A_3)\cdot Pr(A_3)$ which equals $(1)\cdot\left(\frac{1}{6}\right) = \frac{1}{6}$. The denominator is as in (i) and (ii). So the

result is that

$$\Pr(A_3 \mid B) = \frac{\left(\frac{1}{6}\right)}{\left(\frac{3}{6}\right)} = \frac{1}{3}.$$

Notice that the denominator is the same in every case.

Let's have another look at Bayes' Theorem. The numerator is the probability of $A_i \cap B$, and the denominator is just the probability of B.

$$\Pr(A_i \mid B) = \frac{\overbrace{\Pr(B \mid A_i) \cdot \Pr(A_i)}^{\boxed{\Pr(A_i \cap B)}}}{\underbrace{\sum_{j=1}^{n} \Pr(B \mid A_j) \cdot \Pr(A_j)}_{\boxed{\Pr(B)}}}$$

Notice that what Bayes' Theorem enables us to do is to <u>reverse the order</u> in conditional probabilities. We use $Pr(B|A_1)$, $Pr(B|A_2)$, ..., $Pr(B|A_n)$ to help us evaluate $Pr(A_1|B)$, $Pr(A_2|B)$ and so on.

<u>E.g.</u> A test for a disease correctly detects that a person has the diease with probability .9, and correctly detects that a person doesn't have the disease with probability .9. If only 1% of the population has the disease, what is the probability that a person shown by the test as having the disease actually does have it?

Sol$^{\underline{n}}$

Let Y_d mean "yes, has the disease",

N_d mean "no, doesn't have the disease",

Y_t mean "yes, tests positive" and

N_t mean "no, tests negative".

Then
$$Pr(Y_d) = .01 \text{ and } Pr(N_d) = .99.$$

Also, $Pr(Y_t \mid Y_d) = .9 = Pr(N_t \mid N_d)$.

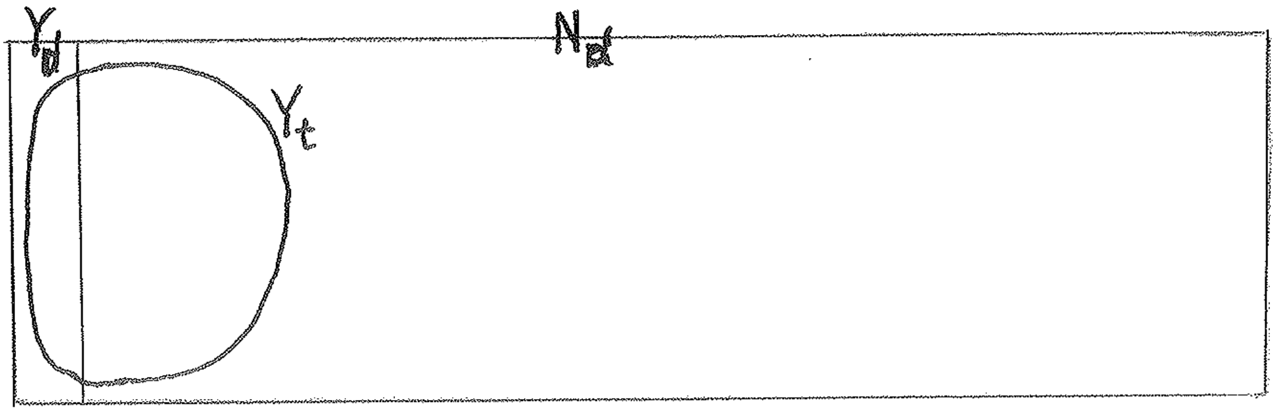
So
$$Pr(Y_d \mid Y_t) = \frac{Pr(Y_t \mid Y_d) \cdot Pr(Y_d)}{Pr(Y_t \mid Y_d) \cdot Pr(Y_d) + Pr(Y_t \mid N_d) \cdot Pr(N_d)}$$

and since $Pr(Y_t \mid N_d) = 1 - Pr(N_t \mid N_d) = .1$

we have
$$Pr(Y_d \mid Y_t) = \frac{(.9)(.01)}{(.9)(.01) + (.1)(.99)}$$
$$= \frac{.009}{.009 + .099}$$
$$= \frac{.009}{.108} = \frac{9}{108} = \frac{1}{12}.$$

So only $1/12$ of those tested positive actually has the disease.

How do we explain this result? The diagram below may help. Although test outcome Y_t occupies 90% of Y_d and only 10% of N_d, because Y_d (the sub-population of people with the disease) is so small it only occupies $1/12$ of Y_t.



(Not quite to scale)

So Bayes' Theorem enables us to reach a conclusion which, although possibly surprising, is correct.